Celal GENÇOĞLU – Orcid ID: 0000-0001-5568-1331
Hikmet GÜMÜŞ – Orcid ID: 0000-0001-7671-4868
Fırat ÖZDALYAN – Orcid ID: 0000-0003-3577-0235
Sercin KOSOVA – Orcid ID: 0000-0003-1005-3387
Egemen MANCI – Orcid ID: 0000-0001-8965-4884

Geliş Tarihi (Received): 23.01.2019
Kabul Tarihi (Accepted): 08.05.2019
SPORMETRE, 2019,17(3), 175-185
DOI: 10.33689/spormetre.516683

FUTBOLDA LİG SIRALAMASINDA PAGERANK ALGORİTMASI YAKLAŞIMI

Celal GENÇOĞLU[1], Hikmet GÜMÜŞ[2], Fırat ÖZDALYAN[3], Sercin KOSOVA[4], Egemen MANCI[5]
[1,2,3,4,5] Dokuz Eylül Üniversitesi, Spor Bilimleri Fakültesi, İzmir

Öz: Bu araştırmanın amacı PageRank ağırlıklı listeleme algoritması kullanarak bir futbol ligi sezonunda takımları kazanma-kaybetme, ofansif ve defansif performanslarına göre sıralamaktır. Türkiye Futbol Süper Ligi 2017-18 sezonuna katılan 18 takımın aralarında çift devreli lig usulüne göre oynadığı 306 maç sonuçları, atılan ve yenilen goller ve puan tablosu Türkiye Futbol Federasyonu internet sitesinden alınmıştır. Veriler Matlab yazılımı ile işlenerek hesaplamalar yapılmıştır. Takımların ligde kazanan, yenilen ve berabere kalmaları durumunda sırasıyla 3 puan, 0 puan ve 1'er puan aldıkları resmi statüde elde ettikleri puanlarla oluşan lig sıralaması ve PageRankdirekt, PageRankofans, PageRankdefans ve PageRankfark değerleri arasındaki ilişki Spearman 's rho nonparametrik korelasyon ile değerlendirilmiştir. PageRank algoritmasının kazanma-kaybetme, ofansif ve defansif performansa göre uygulandığı ağırlıklı liste sıralamasının karşılaştırıldığı bu çalışmada takımların başarı sıralamalarının değiştiği görülmüştür. $PageRank_{direkt}$ ve $PageRank_{fark}$ sıralamasında ilk 4 sırada yer alan en başarılı takımlar aynı kalırken resmi puanlamadakinden farklı sıralandıkları tespit edilmiştir. Benzer olarak son üçte yer aldıkları için bir alt lige düşen takımlar da hem $PageRank_{direkt}$ hem de $PageRank_{fark}$ sıralamasında resmi statüden farklı çıkmıştır. Bu çalışma sonuçları takımların güçlü veya zayıf rakiplerinden aldıkları puanların farklı değerde olduğunu ve ağırlığına göre puan almalarının daha doğru olduğunu ortaya koymaktadır. Ayrıca kazanılan maçlardaki skor farkı veya atılan, yenilen goller de göz önüne alınarak başarı sıralaması yapılırsa daha objektif bir performans değerlendirmesi olacağı gösterilmiştir. Dahası yarışmalarda dağıtılacak puanı belirleyecek bir matematiksel model geliştirilebilir. Bu çalışmanın bulguları takımların sıralanmasında kullanılacak alternatif yöntemleri tartışmak adına önemli farklar ortaya koymuştur. Yeni sıralama yaklaşımının daha adil ve kaliteli bir lig sıralaması yapmasının mümkün olduğu test edilmiş ve düz puan sistemine göre objektif avantajları öne çıkmıştır. İleri araştırmalarda PageRank algoritması modifiye edilerek farklı niteliklerin değerlendirmeye alındığı bir yöntem futbol branşı için önerilebilir.

Anahtar Kelimeler: PageRank, Sıralama algoritmaları, Sportif sıralama, Futbol,

PAGERANK ALGORITHM APPROACH TO RANKING SOCCER TEAMS

Abstract: The aim of this study compares to the traditional ranking system and Pagerank algorithm rankings for a season of Turkish Super Soccer League. Three hundred six matches results including the 2017-18 season of Turkish Super Soccer League collected from the Turkish Soccer Federation website and transferred to Matlab software. Relationship between the traditional ranking system and PageRank algorithms such as win-lose, difference, offense and defense values were tested with Spearman nonparametric correlation. We evaluate the PageRank algorithm ranked the teams significantly different from the traditional ranking system. PageRankdirect algorithm ranked league's first four same teams but changed the placements. This reveals that the medal table and cup holders can be changed. Also, the algorithm yielded standings of last three teams mostly same one exception. Moreover, PageRankdifference, PageRankoffence and PageRankdefence values showed significantly different rankings teams when compared to the official point system. In this study, the PageRank algorithm presented a novel approach to ranking soccer teams. This method has limitations likewise four different components of evaluation reveals four separate outputs. Further researches needed to focus on only one PageRank algorithm for these performance indicators such as win-lose, difference, offensive, and defensive values. The conclusion we do not conceive the traditional substituting system within the PageRank algorithm for league standings, but it could be regarded as a method of assessment team market values and distributions of income.

Key Words: PageRank, Ranking algortihms, Sports Ranking, Soccer

Celal GENÇOĞLU – Orcid ID: 0000-0001-5568-1331
Hikmet GÜMÜŞ – Orcid ID: 0000-0001-7671-4868
Fırat ÖZDALYAN – Orcid ID: 0000-0003-3577-0235
Sercin KOSOVA – Orcid ID: 0000-0003-1005-3387
Egemen MANCI – Orcid ID: 0000-0001-8965-4884

GİRİŞ

Tüm spor dallarında takımların ve sporcuların performans değerlendirmesi sonucunda sıralamalar yapılmaktadır. Bu sıralama sonucu antrenörler için oyuncu değerlendirme imkanı verdiği gibi, takımlar için sponsorluk ve diğer mali anlaşmalar sağlaması, yayın gelirleri, bilet, forma ve ürün satışları açısından kritik öneme sahiptir. Başarılı takımlar, diğer bir deyişle daha üst sırada yer alan "Şampiyon" unvanı alan takımlar daha popüler olarak daha fazla taraftar kazanmaktadır.

Futbolda takımlar lig veya grup maçlarında kazandıkları galibiyetler ve beraberliklerden aldıkları puanların toplamına göre sıralanmaktadır. Bu resmi statüye göre takımlar kazanma, beraberlik ve yenilgi için sırasıyla üç, bir ve sıfır puan kazanırlar. Ancak takımların birbirlerine karşı aldıkları sonuçların diğer takımlara olan etkisi bu sistemle göz ardı edilmektedir. İyi bir sıralama yöntemi ile takımların sadece aldığı galibiyet değil yendikleri veya yenildikleri takımın gücü de önem kazanmaktadır. Bazı spor dalları veya spora ait uluslararası federasyonlar kulüp takımlarını, milli takımları ve sporcuları sıralamak için farklı yöntemler kullanmaktadır. Bunlardan Kazanma Yüzdesi (Winning Percentage: WP) takımları basitçe kazandıkları maça göre sıralar. Puanlama Yüzde Endeksi (The Rating Percentage Index: RPI) kazanma yüzdesi temel almaktadır ve RPI skoru kendisinin, rakiplerinin ve onların rakiplerinin yüzdesinin ağırlıklı ortalamasıdır (Barrow ve ark., 2013).

Futbolda sıralama yöntemi olarak takımların kime karsı kazandıklarının ya da kaybettiklerinin de değerlendirilmesi için PageRank ağırlıklı sıralama algoritması kullanılabilir. PageRank algoritması (Page ve ark., 1999) Google şirketinin kurucuları Larry Page ve Sergey Brin tarafından geliştirilen ve arama sonuçlarının daha anlamlı şekilde sıralanması için kullanılan bir algoritmadır. Geliştirilen algoritma takım sporlarında bireysel sporcu performansını göstermek için oyunun merkezinde yer alan kilit roldeki oyuncuyu göstermek için kullanılmıştır. Buna göre algoritma yöntemi sporcuların arasındaki bağlantıyı ve bu bağlantıda önemli oyuncuyu bularak başarılı, önem derecesi üst sıradaki sporcuları bulmaya çalışmıştır (Brandt ve ark., 2015; Brown ve ark., 2017). Yine, bireysel performansın objektif değerlendirilebilmesi için Atletizm ve tenis sporcularının sıralaması bu yöntemin uygulaması ile karşılaştırılmıştır (Beggs ve ark., 2017; London ve ark., 2014). Bunun yanında söz konusu yöntem kolej sporları antrenörlerini sıralamak için de kullanılmıştır (Hu ve ark., 2015). Spor alanında kullanımı giderek yaygınlaşan PageRank ağırlıklı listeleme sıralaması basketbol (Barrow ve ark., 2013), beyzbol (Govan ve ark.,2006), Amerikan futbolu (Zack ve ark., 2012) gibi takım sporlarında sıralama yönteminin araştırılmasında kullanılmıştır. Literatürde futbolda sıralamanın yapılmasında PageRank algoritması kullanılan bilinen sadece bir araştırma vardır (Lazova ve ark., 2015). Bu çalışmada Ulusal takımların sıralaması ile Uluslararası Futbol Federasyonları Birliği (FIFA) takımlar sıralaması karşılaştırılmıştır. Ayrıca sadece kazanma-kaybetme üzerinden PageRank değerleri elde edilmiştir. Bu araştırmada bir futbol ligi sezonunda takımların kazanma-kaybetme, ofansif ve defansif performanslarına göre PageRank algoritmasına göre sıralanması araştırılmıştır. $PageRank_{ofans}$ ve $PageRank_{defans}$ sonuçları sıralamanın resmi sıralamaya göre hem ilk dört takımın hem de son üç takımın yerlerinin değiştiğini göstermektedir.

MATERYAL VE METOT

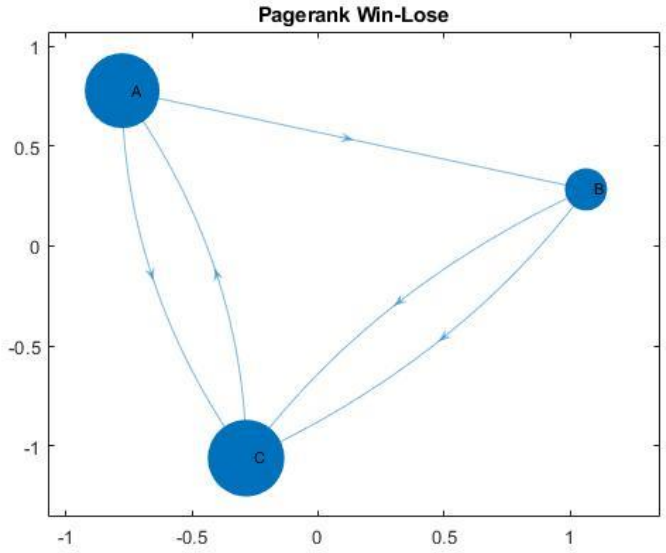
Celal GENÇOĞLU – Orcid ID: 0000-0001-5568-1331
Hikmet GÜMÜŞ – Orcid ID: 0000-0001-7671-4868
Fırat ÖZDALYAN – Orcid ID: 0000-0003-3577-0235
Sercin KOSOVA – Orcid ID: 0000-0003-1005-3387
Egemen MANCI – Orcid ID: 0000-0001-8965-4884

Veri

Türkiye Futbol Süper Ligi 2017-18 sezonuna katılan 18 takımın aralarında çift devreli lig usulüne göre oynadığı 306 maç sonuçları, atılan ve yenilen goller ve puan tablosu Türkiye Futbol Federasyonu internet sitesinden alınmıştır. Veriler Matlab yazılımı ile işlenerek hesaplamalar yapılmıştır (MATLAB and Statistics Toolbox Release 2018, The MathWorks, Inc., Natick, Massachusetts, United States).

PageRank Algoritması

Algoritmanın takımların kazanma ve kaybetmelerde kendilerine başlangıçta atanan belirli bir değerin bir kısmını karşı tarafa vermeleri temeline dayanmaktadır. Algoritma birden çok tez tekrarlanarak (iterasyon) sonuçların değişmeyeceği bir noktaya ulaşması sağlanır. Algoritma başladığında eğer önceden tanımlanmış bir ilk değer yok ise o zaman her takıma 1/takım sayısı kadar puan verilir. Bir sonraki tekrarda takım, yenildiği takıma bu değerin bir kısmını aktarır (Şekil 1). Aktardığı değer takımın bu turda sahip olduğu değerin takımın yenilme sayısına bölünmesiyle elde edilir. Bu şekilde bir takım ne kadar çok takıma yenildiyse onlara aktaracağı değerde düşük olacaktır.



Şekil 1. PageRank algoritmasında takımların (A, B ve C) kazanan takımlara link (değer) vermesi

Örnek tabloda (Tablo 1) takımların kazanma ve kaybetme sayıları bulunmaktadır. Grafik üzerinde ise takımların yaptığı maçlar görülmektedir. Buradan anlaşılacağı üzere, 3 galibiyet ve 1 yenilgi alan C takımı en yüksek PR değerini almaktadır. Bu noktada dikkati çeken kısım ise aynı sayıda galibiyet ve yenilgisi olan iki takımın nasıl sıralanacağıdır. Her iki takımda bir galibiyet almasına rağmen A takımı birinci sırada olan C takımını yendiği için ondan daha fazla değer alarak ikinci sırada yer almaktadır. Böylece elde ettikleri sonuçları birbirileriyle ilişkilendirerek ağırlıklı listeleme ile sıralama yapılmıştır.

Tablo 1. Takımların (A, B ve C) maçlarda aldıkları galibiyet e yenilgiler ve PageRank değerleri

Sıralama	Galibiyet	Yenilgi	Takım
0.39738	3	1	'C'
0.38777	1	2	'A'
0.21485	1	2	'B'

PageRank Hesaplamaları

Direkt; Kazanma-Kaybetme, Ofansif; Atılan gol, Defansif; Yenen gol, Fark; Kazanılan veya kaybedilen maçtaki skor farkı olmak üzere 4 farklı PageRank değeri hesaplanmıştır. $PageRank_{direkt}$ hesaplamasında iterasyonlarda takımlar değerlerini kaybettikleri takıma aktarmaktadır. $PageRank_{direkt}$ hesaplanırken beraberlikler herhangi bir takım yönüne doğru olmadığından değerler değişimi olamayacağı için hesaplamaya alınmamıştır. Burada takımların kazandığı maçlar kaybeden takımdan kendilerine verilen bir link olarak değerlendirilmiştir. $PageRank_{ofans}$ hesaplamasında ise atılan goller link olarak gösterilmiş ve değer gol yiyen takımdan gol atan takıma aktarılmıştır. $PageRank_{defans}$ hesaplamasında ise yenen goller bir link olarak gösterilmiştir. Bu durumda ise değer gol yiyen takıma doğru aktarıldığından sıralamada en düşük değere sahip takım daha iyi olarak değerlendirilmiştir. $PageRank_{ofans}$ ve $PageRank_{defans}$ değerleri takımların birbirlerine karşı kaydettikleri golleri dikkate aldığı için beraberlikle sonuçlanan maçlar da hesaplamalara dahil edilmiştir.

PageRank hesaplamalarında iterasyon 500 kez olana kadar tekrarlanmıştır. Algoritmada tüm PR değerlerinin toplamı "1" olarak çalışılmıştır. Takımların PageRank değerlerinin hassas olarak karşılaştırılabilmesi için işlemler nokta sonrası 6 basamaklı floating point veri tipinde alınmıştır. Yapılan iterasyonlar sırasında takımların sahip oldukları değerin ne kadarını aktaracakları "Damping faktörü" ile belirlenmiş ve bu çalışmada 0.85 olarak alınmıştır. Bu veriler kullanılarak takımlar arasındaki ilişkileri gösteren bir network grafiği oluşturularak görselleştirilmiştir.

Korelasyon

Takımların ligde kazanan, yenilen ve berabere kalmaları durumunda sırasıyla 3 puan, 0 puan ve 1'er puan aldıkları resmi statüde elde ettikleri puanlarla oluşan lig sıralaması ve $PageRank_{direkt}$, $PageRank_{ofans}$, $PageRank_{defans}$ ve $PageRank_{fark}$ değerleri arasındaki ilişki Spearman 's rho nonparametrik korelasyon ile değerlendirilmiştir. Hopkins'e göre korelasyonun önemi skalası ($r<0.1$,önemsiz; 0.1-0.3, düşük; 0.3-0.5, orta; 0.5-0.7, yüksek; 0.7-0.9, çok yüksek; >0.9, neredeyse mükemmel ve 1 mükemmel) kullanılmıştır (Hopkins ve ark., 2008).

BULGULAR

Resmi statüye göre puanlama sıralaması ile $PageRank_{direkt}$, $PageRank_{ofans}$, $PageRank_{defans}$ ve $PageRank_{fark}$ değerlerine göre oluşturulan ağırlıklı liste sıralamasında takımların yerleri farklı bulunmuştur. Biri geleneksel yöntem ve dört farklı PageRank algoritmasına göre olmak üzere tüm puan, PageRank değerleri ve sıralamalar Tablo 3 de sunulmuştur.

$PageRank_{direkt}$ değerinin takımlar arasındaki network ilişkisini gösteren grafik oluşturulmuş ve renklendirilerek görselleştirilmiştir (Şekil 2).

Resmi puanlama sistemi ve tüm PageRank değerleri arasında anlamlı pozitif yüksek korelasyon bulunmuştur (Tablo 2). Yalnızca $PageRank_{ofans}$ ve $PageRank_{defans}$ değerleri orta düzeyde anlamlı pozitif korelasyon göstermiştir.

Celal GENÇOĞLU – Orcid ID: 0000-0001-5568-1331
Hikmet GÜMÜŞ – Orcid ID: 0000-0001-7671-4868
Fırat ÖZDALYAN – Orcid ID: 0000-0003-3577-0235
Sercin KOSOVA – Orcid ID: 0000-0003-1005-3387
Egemen MANCI – Orcid ID: 0000-0001-8965-4884

Tablo 2. Resmi puanlama sistemi ve PageRank algoritmaları arasındaki ilişki

		Resmi puanlama	**PageRank$_{direkt}$**	**PageRank$_{fark}$**	**PageRank$_{ofans}$**
PageRank$_{direkt}$	*r*	,998**			
	p	,000			
PageRank$_{fark}$	*r*	,998**	1,000**		
	p	,000			
PageRank$_{ofans}$	*r*	,780**	,787**	,787**	
	p	,000	,000	,000	
PageRank$_{defans}$	*r*	-,668**	-,674**	-,674**	-,521*
	p	,002	,002	,002	,027

**Korelasyon anlamlılık seviyesi 0.01
*Korelasyon anlamlılık seviyesi 0.05

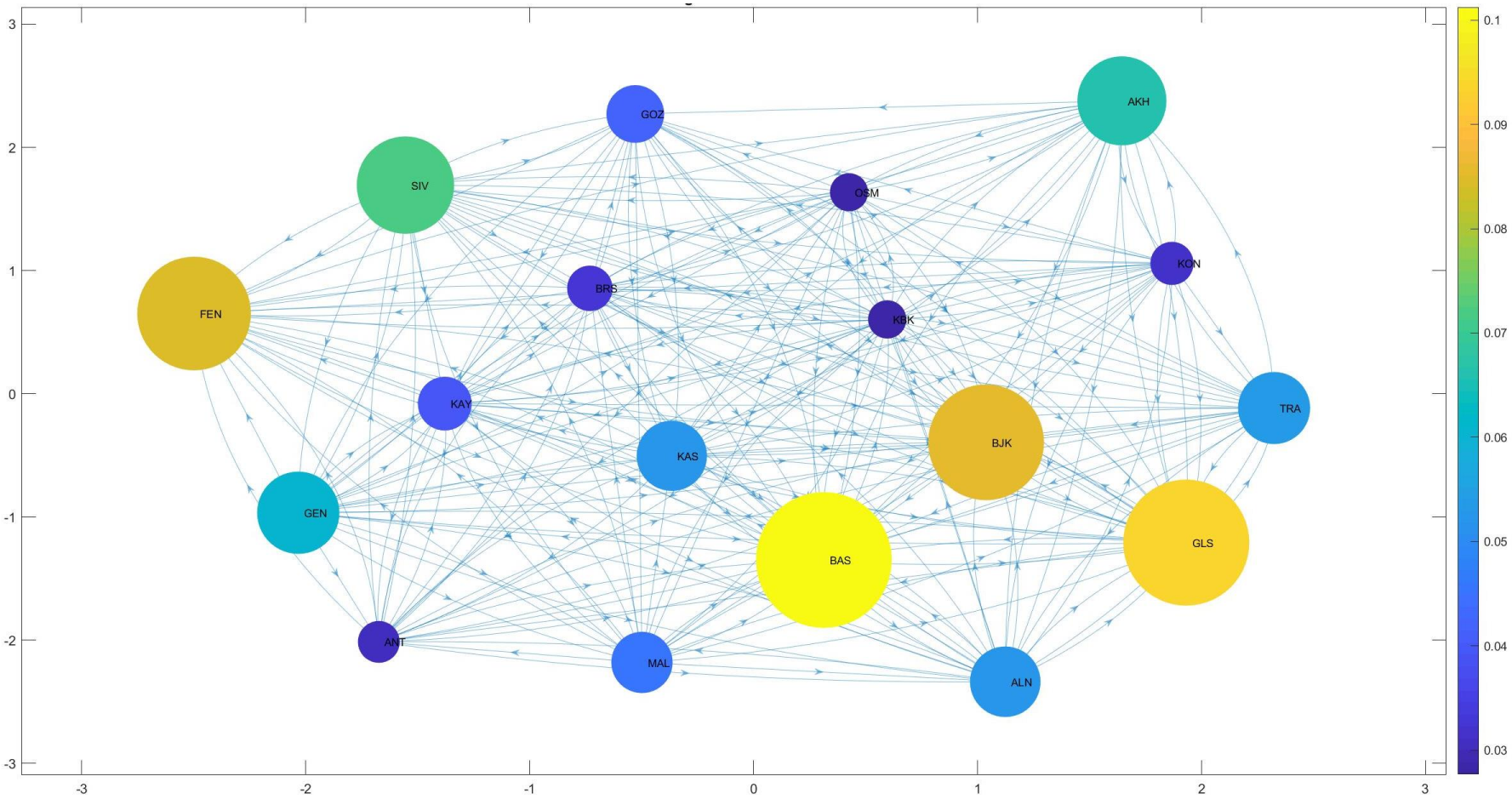
Celal GENÇOĞLU – Orcid ID: 0000-0001-5568-1331
Hikmet GÜMÜŞ – Orcid ID: 0000-0001-7671-4868
Fırat ÖZDALYAN – Orcid ID: 0000-0003-3577-0235
Sercin KOSOVA – Orcid ID: 0000-0003-1005-3387
Egemen MANCI – Orcid ID: 0000-0001-8965-4884

Tablo 3 Resmi Sistem Puanları ve PageRank Değerleri Sıralamaları Karşılaştırılması

Resmi Sistem Puanları ve Sıralaması						$PageRank_{direkt}$ Değerleri ve Sıralaması				$PageRank_{fark}$ Değerleri ve Sıralaması				$PageRank_{ofans}$ Değerleri ve Sıralaması				$PageRank_{defans}$ Değerleri ve Sıralaması			
Sıra	Takım	Galibiyet	Beraberlik	Yenilgi	Puan	Takım	Galibiyet	Yenilgi	$PageRank_{direkt}$ Değeri	Takım	Galibiyetteki Skor Farkı	Yenilgideki Skor Farkı	$PageRank_{fark}$ Değeri	Takım	Atılan Gol Sayısı	Yenilen Gol Sayısı	$PageRank_{ofans}$ Değeri	Takım	Yenilen Gol Sayısı	Atılan Gol Sayısı	$PageRank_{defans}$ Değeri
1	GLS	24	3	7	75	BAS	22	6	0,10116846	GLS	54	12	0,114297909	FEN	78	36	0,076415356	KBK	86	20	0,083273963
2	FEN	21	9	4	72	GLS	24	7	0,093871301	BJK	45	6	0,112107446	BJK	69	30	0,071315765	ALN	59	55	0,064852712
3	BAS	22	6	6	72	BJK	21	5	0,086138442	BAS	39	11	0,106722611	GLS	75	33	0,070592447	ANT	59	40	0,063740929
4	BJK	21	8	5	71	FEN	21	4	0,084538321	FEN	47	5	0,091547439	TRA	63	51	0,067997396	KAY	55	44	0,062112473
5	TRA	15	10	9	55	SIV	14	13	0,072262924	AKH	19	28	0,063092113	BAS	62	34	0,06728781	GEN	54	37	0,061112423
6	GOZ	13	10	11	49	AKH	11	14	0,065974571	ALN	21	25	0,060304677	KAS	57	58	0,065874803	KAS	58	57	0,059672952
7	SIV	14	7	13	49	GEN	8	17	0,060898729	SIV	20	28	0,058877792	ALN	55	59	0,062948638	OSM	60	49	0,059149878
8	KAS	13	7	14	46	TRA	15	9	0,053247437	GEN	14	31	0,054187389	OSM	49	60	0,054034335	AKH	53	44	0,058803139
9	KAY	12	8	14	44	ALN	11	16	0,052104124	TRA	29	17	0,048369173	SIV	45	53	0,053913805	TRA	51	63	0,057137167
10	MAL	11	10	13	43	KAS	13	14	0,051884975	KAS	22	23	0,046185781	GOZ	49	50	0,052820503	MAL	45	38	0,054967334
11	AKH	11	9	14	42	MAL	11	13	0,045123706	KON	19	23	0,034779411	AKH	44	53	0,052345247	SIV	53	45	0,054766099
12	ALN	11	7	16	40	GOZ	13	11	0,042463358	MAL	26	23	0,032845323	KAY	44	55	0,050905131	BRS	48	43	0,054023198
13	BRS	11	6	17	39	KAY	12	14	0,039536823	OSM	20	31	0,032317234	GEN	37	54	0,047233143	GOZ	50	49	0,051274196
14	ANT	10	8	16	38	BRS	11	17	0,033224102	BRS	22	27	0,032139243	BRS	43	48	0,046365217	BAS	34	62	0,046825691
15	KON	9	9	16	36	KON	9	16	0,031498638	KAY	16	27	0,029475715	KON	38	42	0,046011881	KON	42	38	0,04495166
16	OSM	8	9	17	33	ANT	10	16	0,030531662	GOZ	20	21	0,028908432	ANT	40	59	0,04433619	GLS	33	75	0,044528299
17	GEN	8	9	17	33	KBK	3	28	0,027837511	KBK	4	70	0,027799885	MAL	38	45	0,040481258	FEN	36	78	0,042805063
18	KBK	3	3	28	12	OSM	8	17	0,027694915	ANT	15	34	0,026042401	KBK	20	86	0,029121075	BJK	30	69	0,036002823

GLS: Galatasaray, FEN: Fenerbahçe, BAS: Medipol Başakşehir, BJK: Beşiktaş, TRA: Trabzonspor, GOZ: Göztepe, SIV: Sivasspor, KAS: Kasımpaşa, KAY: Kayserispor, MAL: Yeni Malatyaspor, AKH: Akhisarspor, ALN: Aytemiz Alanyaspor, BRS: Bursaspor, ANT: Antalyaspor, KON: Atiker Konyaspor, OSM: Osmanlıspor FK, GEN: Gençlerbirliği, KBK: Karabükspor

Celal GENÇOĞLU – Orcid ID: 0000-0001-5568-1331
Hikmet GÜMÜŞ – Orcid ID: 0000-0001-7671-4868
Fırat ÖZDALYAN – Orcid ID: 0000-0003-3577-0235
Sercin KOSOVA – Orcid ID: 0000-0003-1005-3387
Egemen MANCI – Orcid ID: 0000-0001-8965-4884



Şekil 2. PageRank$_{direkt}$ network grafiği

Celal GENÇOĞLU – Orcid ID: 0000-0001-5568-1331
Hikmet GÜMÜŞ – Orcid ID: 0000-0001-7671-4868
Fırat ÖZDALYAN – Orcid ID: 0000-0003-3577-0235
Sercin KOSOVA – Orcid ID: 0000-0003-1005-3387
Egemen MANCI – Orcid ID: 0000-0001-8965-4884

TARTIŞMA

Futbolda sezon sonu resmi lig sıralaması ile PageRank algoritmasının kazanma-kaybetme, ofansif ve defansif performansa göre uygulandığı ağırlıklı liste sıralamasının karşılaştırıldığı bu çalışmada takımların başarı sıralamalarının değiştiği görülmüştür. Bu çalışmayı diğerlerinden ayıran en önemli özelliği kazanma-kaybetme, skor farkı, atılan ve yenilen gol için dört farklı veriden PageRank sıralaması yapmasıdır. $PageRank_{direkt}$ ve $PageRank_{fark}$ sıralamasında ilk 4 sırada yer alan en başarılı takımlar aynı kalırken resmi puanlamadakinden farklı sıralandıkları tespit edilmiştir. Benzer olarak son üçte yer aldıkları için bir alt lige düşen takımlar da hem $PageRank_{direkt}$ hem de $PageRank_{fark}$ sıralamasında resmi statüden farklı çıkmıştır. Bunun dışında en çok galibiyeti alıp en çok puan toplayan takım $PageRank_{direkt}$ değerine göre ligi ilk sırada bitirip şampiyon unvanını alamamıştır. Bu çalışma sonuçları takımların güçlü veya zayıf rakiplerinden aldıkları puanların farklı değerde olduğunu ve ağırlığına göre puan almalarının daha doğru olduğunu ortaya koymaktadır. PageRank algoritmasıyla takımların sıralamasının resmi sıralamaya göre değişmesine rağmen geleneksel puan toplama sistemiyle tüm PageRank değerleri arasında anlamlı pozitif korelasyon bulunmuştur. Bulunan ilişki PageRankdirekt ve PageRank fark ile puan toplama sistemi arasında çok yüksek anlamlı korelasyon görülürken PageRankofans ve PageRankdefans arasında orta düzeyde korelasyon görülmüştür. Buna göre kazanılan maçlardaki skor farkı veya atılan, yenilen goller de göz önüne alınarak başarı sıralaması yapılırsa daha objektif bir performans değerlendirmesi olacağı gösterilmiştir. Ayrıca atılan ve yenen gollerle takımların hücum ve savunma performanslarının değerlendirilmesi ve sıralanması resmi sisteme göre oldukça ayırıcı bir etkendir.

Aslen, Amerikan Kolej Futbolu Ligi takımları sıralamak için Coley matriks yöntemi (Govan ve ark., 2008) ve Massey asgari kareler sistemi (Massey, 1997) gibi doğrusal matematik tabanlı sıralama yöntemlerini kullanmaktadır. Çok sayıdaki takımın konferanslara ayrıldığı ve konferans içi ve dışı maç sayısının farklı olması gibi nedenlerin yanında kazanma yüzdesinin tek başına takımları sıralamakta yeterli olmadığı düşünülmektedir (Langville ve ark., 2012). Bu çalışmada belirlenen sıralama farkı Lazova ve ark. (2015) yaptığı çalışmada görülen FIFA sıralaması ile PageRank sıralamasında ilk 20 takım için sıralamanın farklı olduğunu gösterilmiştir. Ancak FIFA organizasyonlarında tüm zamanlarda oynanan maçlara ait bir data setiyle çalışmışlardır. Buna göre tüm 210 ülke takımı ve grup maçları da dahil her müsabakanın algoritmaya dahil edilmiş olması ilk 20 takımın dizilmesinde bu çalışmayla karşılaştırmayı mümkün kılmamaktadır. Ayrıca, ülke milli takımlarını sıralarken FIFA resmi puan sistemi yerine Elo's yöntemini kullanmaktadır. Söz konusu yöntem de takımların geçmiş performanslarına göre yapılan tahmin ve gerçekleşen sonuç arasındaki farka göre sıralama yapmaktadır.

Başarı diziliminde skor üretme becerisi dikkate alındığında, $PageRank_{ofans}$ ve $PageRank_{defans}$ sonuçları sıralamanın resmi sıralamaya göre hem ilk dört takımın hem de son üç takımın yerlerinin değiştiğini göstermektedir. Buna göre takımların güçlü rakiplere karşı gol atabilmeleri ya da daha zayıf rakiplerinden az gol yemeleri ağırlıklı listelemede daha iyi yerlere sahip olmalarını sağlamaktadır. Buna benzer bir yaklaşımı uygulayan voleybol branşı, takım sıralama sistemi olarak futbolda uygulanan statüye

çok benzer 3 puanlı sistemi kullanır ancak farklı olarak takımların kazandıkları setlere de önem vererek yenilmesine rağmen 2 set kazanan takıma da puan verir. Yine de ulusal ve uluslararası federasyon müsabaka talimatları (2018 Aralık 14) voleybolda set içinde alınan sayılar ve kimin kimi yendiği göz ardı edilmektedir. Oysa, $PafeRank_{ofans}$ ve $PageRank_{defans}$ algoritma sıralaması ile takımların sadece skor üretme veya skora izin vermeme becerisi değil bu performansı gösterdikleri takımın gücü de değerlendirmeye alınmıştır. Futbolda resmi puanlama statüsünde olmayan bu uygulama $PageRank_{ofans}$ ve $PageRank_{defans}$ değerleri ile gösterilmiştir. Futbolda izlenirliği artıran en önemli etkenlerden biri maçlarda çok gol atılmasıdır. Örneğin, dünya futbolunda profesyonel liglerde en yüksek maç başına skor 3.97 gol ortalamayla Almanya birinci futbol ligi Bundeliga'dır. Yine, Bundesliga dünyada maç başına ortalama seyirci sayısı (2016-17 sezon ort 41511; 2017-18 sezonu 43879) en yüksek lig olarak öne çıkmaktadır. Aynı zamanda 2016-17 sezonunda ligin televizyon izlenme sayısı yaklaşık üç milyar kişidir (www.statista.com). Resmi puan sisteminde olmayıp bu çalışmada önerilen algoritma gibi maçlarda atılan gol ve atıldığı rakibin gücü değerlendirmeye alınsaydı takımların uyguladığı stratejiler de değişebilirdi. Kimi güçsüz takımlar güçlü takımlara karşı oynadıkları kapalı defans kurgusu yerine gol atmaya öncelik verebilirlerdi ki bu da oyuna seyir zevki açısından önemli bir değer katabilirdi. Literatüre göre sadece Govan ve ark (2009) takımları hücum ve savunma performanslarına göre sıralama yaklaşımını araştırmışlardır. Ancak çalışmalarında doğrusal cebir yöntemi ve web sitelerinin sıralanmasında kullanılan benzer bir algoritma olan Kleinberg's Hyperlink-Induced Topic Search kullanmışlardır. Bu araştırmada ise hücum-savunma performansı PageRank algoritmasıyla ağırlıklı olarak değerlendirilmiştir. Aslında futbolda takımların defansif ve ofansif yönleri tartışılan bir konudur. Sıklıkla değerlendirilen hücum ve savunma kapasitelerinin objektif bir yöntemle belirlenmesi sıralamaya etki etmemesine rağmen oldukça önemlidir. Diğer yandan, kazanılan ya da kaybedilen maçlarda skor üstünlüğüne de önem verildiği takdirde ($PageRank_{fark}$ değerlerine göre) sıralama son sıralardaki takımların farklı yerlerde sezonu bitirebileceğini belirlemiştir. Buna göre, kazanılan veya kaybedilen maçların skor açısından ne kadar dengeli ya da dengesiz olduğunu gösteren bu dereceleme maç skorunun, maçta sonucunu yansıtmasa bile tüm sezondaki maçlara ve rakiplere göre değerlendirilebileceğini göstermektedir. Literatürde oldukça sınırlı sayıda çalışma skor farkını gözeterek takımları değerlendirmiştir. Barrow ve ark. (2013) çalışmalarında skor farkını baz alan algoritma değerleri hesaplamışlardır. Skor farkı verisinin sıralamayı belirlemekte kazanma-kaybetme verisine göre daha doğru bir belirleyici olduğunu göstermişlerdir. Ayrıca, Amerikan kolej basketbol takımlarının turnuva kazanma şanslarına göre matematiksel yöntemler kullanılarak sıralandığı çalışmalar bulunmaktadır (Jacobson H., 2009). Bu sıralama sistemini bahis ve olasılık amaçlı araştırsa da yöntem kazanma-kaybetme kayıtları yanında kazanma farklarını da dikkate almaktadır. Fakat, turnuvada eleme maçlarının da etkisiyle daha üst sıraların doğru sıralamasını tahmin etme açısından kararsız sonuçlar bulunmuştur. Sonuçta, maç sonucundaki üstünlük farkı, takımlarda farklı önde veya geride oldukları karşılaşmalarda oyundan kopmalara neden olabildiğinden bu çalışmadaki $PageRank_{fark}$ değerine göre sıralama oyun temposunu gözeten, maçlarda kopmaları engelleyen, daha çok gol atılmasını sağlayan bir etken olarak öne çıkmaktadır.

Bu çalışmada kullanılan dört farklı PageRank algoritması her biri için farklı sonuçlar ortaya koymuştur ve resmi sıralamayla karşılaştırılmıştır. Kazanma-kaybetmenin, skor farkının, atılan ve yenen gollerin tek bir PageRank algoritmasında değerlendirilmemesi araştırmanın kısıtlılığıdır. İleri araştırmalarda PageRank algoritması modifiye edilerek farklı niteliklerin değerlendirmeye alındığı bir yöntem futbol branşı için önerilebilir. Dahası yarışmalarda dağıtılacak puanı belirleyecek bir matematiksel model geliştirilebilir.

SONUÇ

Sporda kayıtlanan veri büyüdükçe, maçlara ait hem takım hem de oyuncu bazında istatistiksel bilgi biriktikçe geriye dönük çıkarım yapmak üzere veri işleme şansı artmaktadır. Bunun en iyi örneği artık Spor biliminde "Büyük Veri" ve "Veri Analizi" gibi kavramların oturmasıdır. Bu çalışmanın bulguları takımların sıralanmasında kullanılacak alternatif yöntemleri tartışmak adına önemli farklar ortaya koymuştur. Yeni sıralama yaklaşımının daha adil ve kaliteli bir lig sıralaması yapmasının mümkün olduğu test edilmiş ve düz puan sistemine göre objektif avantajları öne çıkmıştır. Son olarak resmi sıralamanın belirlendiği eski statülerin yerini alması beklenmese de takımların market değerinin belirlenmesinde ve gelir dağılımında dikkate alınmasında kullanılabileceği düşünülmektedir.

KAYNAKLAR

Barrow D., Drayer I., Elliott P., and et al. (2013). Ranking rankings: an empirical comparison of the predictive power of sports ranking methods. *JQAS*; 9(2): 187–202.

Beggs CB., Shepherd SJ., Emmonds S., Jones B. (2017) A novel application of PageRank and user preference algorithms for assessing the relative performance of track athletes in competition. *PLoS ONE* 12(6): e0178458. https://doi.org/10.1371/journal.pone.0178458.

Brandt M. and Brefeld U. (2015). Graph-based Approaches for Analyzing Team Interaction on the Example of Soccer, Proceedings of the ECML/PKDD Workshop on Machine Learning and Data Mining for Sports Analytics.

Brown Shael. (2017). A PageRank Model for Player Performance Assessment in Basketball, Soccer and Hockey. arXiv:1704.00583v1 [stat.AP] 31 Mar 2017 .

Federation Internationale De Volleyball Official Rules 2017-2020 https://www.fivb.org/EN/Refereeing-Rules/documents/FIVB-Volleyball_Rules_2017-2020-EN-v06.pdf Son erişim tarihi 14.11.2018 10.20.

Govan AY. (2008). Ranking theory with application to popular sports. Raleigh: North Carolina State University.

Govan AY., Langville AN., and Meyer CD. (2009). Offense-Defense Approach to Ranking Team Sports, *Journal of Quantitative Analysis in Sports*. 5:1; 1-19.

Govan AY., Meyer CD. (2006). Ranking national football league teams using google's pagerank. AA Markov Anniversary Meeting; Charleston: Boson Books.

Hopkins WG. (2013). A New View Of Statistics. Internet Society for Sport Science: http://www.sportsci.org/resource/stats/ 2013.

Hu Z., Zhou J., Zhang M., and Zhao Y. (2015). Methods for ranking college sports coaches based on data envelopment analysis and PageRank. *Expert Systems: The Journal of Knowledge Engineering*, 32, 652–673.

In Statista - The Statistics Portal. Retrieved February 05 2019. Major soccer leagues worldwide ranked by average per game attendance in 2017/18 (in 1,000s) https://www.statista.com/statistics/270301/best-attended-football-stadiums-in-the-world-by-average-attendance-2010.

Jacobson SH. (2009). Seedıng In The Ncaa Men's Basketball Tournament: When Is A Hıgher Seed Better? *The Journal of Gambling Business and Economics*. 3: 2; 63-87.

Langville AN., Meyer CD. (2012). Who's the 1?: The science of rating and ranking. Princeton: Princeton University Press.

Lazova V., Basnarkov L. (2015). PageRank Approach to Ranking National Football Teams arXiv:1503.01331v1 Computer Science, Social and Information Networks [cs.SI] 4 Mar 2015.

London A., Nemeth J., and Nemeth T. (2014). Time-dependent Network Algorithm for Ranking in Sports Acta Cybernetica 21, 495–506.

Massey K. (1997). Statistical models applied to the rating of sports teams. Bluefield College.Page L., Brin S., Motwani R., and Winograd T. (1999). The PageRank Citation Ranking: Bringing Order to the Web. Technical report, Stanford InfoLab Technical Report 1999 – 66.

Türkiye Voleybol Federasyonu Yarışma Talimatı http://www.tvf.org.tr/_dosyalar/Talimatlar/TVF_yarisma_talimati.pdf Son erişim tarihi 14.11.2018 09.58

Zack L., Lamb R., Ball S. (2012). An application of Google's PageRank to NFL rankings. Involve, a *Journal of Mathematics*. 5(4):463±71.